

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2001, 18:30:48 ; Search time 3525.91 Seconds  
(without alignments)  
9884.731 Million cell updates/sec

Title: US-09-117-447-1  
Perfect score: 3687

Sequence: 1 atgataagaaaagctgt.....ttacatctatataagtaa 3687

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est110:\*  
191: gb\_est111:\*  
192: gb\_est112:\*  
193: gb\_est113:\*  
194: gb\_est114:\*  
195: gb\_est115:\*  
196: gb\_est116:\*  
197: gb\_est117:\*  
198: gb\_est118:\*  
199: gb\_est119:\*  
200: gb\_est120:\*  
201: gb\_est121:\*  
202: gb\_est122:\*  
203: gb\_est123:\*  
204: gb\_est124:\*  
205: gb\_est125:\*  
206: gb\_est126:\*  
207: gb\_est127:\*  
208: gb\_est128:\*  
209: gb\_est129:\*  
210: gb\_est130:\*  
211: gb\_est131:\*  
212: gb\_est132:\*  
213: gb\_est133:\*  
214: gb\_est134:\*  
215: gb\_est135:\*  
216: gb\_est136:\*  
217: gb\_est137:\*  
218: gb\_est138:\*  
219: gb\_est139:\*  
220: gb\_est140:\*  
221: gb\_est141:\*  
222: gb\_est142:\*  
223: gb\_est143:\*  
224: gb\_est144:\*  
225: gb\_est145:\*  
226: gb\_est146:\*  
227: gb\_est147:\*  
228: gb\_est148:\*  
229: gb\_est149:\*  
230: gb\_est150:\*  
231: gb\_est151:\*  
232: gb\_est152:\*  
233: gb\_est153:\*  
234: gb\_est154:\*  
235: gb\_est155:\*  
236: gb\_est156:\*  
237: gb\_est157:\*  
238: gb\_est158:\*  
239: gb\_est159:\*  
240: gb\_est160:\*  
241: gb\_est161:\*  
242: gb\_est162:\*  
243: gb\_est163:\*  
244: gb\_est164:\*  
245: gb\_est165:\*  
246: gb\_est166:\*  
247: gb\_est167:\*  
248: gb\_est168:\*  
249: gb\_est169:\*  
250: gb\_est170:\*  
251: gb\_est171:\*  
252: gb\_est172:\*  
253: gb\_est173:\*  
254: gb\_est174:\*  
255: gb\_est175:\*  
256: gb\_est176:\*  
257: gb\_est177:\*  
258: gb\_est178:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72.2	2.0	908	245	AZ548467
2	68.2	1.8	877	245	AZ531291
3	68.2	1.8	912	245	AZ551092
4	66.8	1.8	816	245	AZ535744
5	66	1.8	843	245	AZ551618
6	63.6	1.7	880	245	AZ529191
7	62	1.7	849	245	AZ546009
8	61	1.7	891	248	AZ683582
9	60	1.6	900	245	AZ549980
10	59.8	1.6	997	219	CNS0134P
11	58.8	1.6	494	258	FR0048073
12	58.4	1.6	905	245	AZ550256
13	57.4	1.6	912	245	AZ527724
14	57.4	1.6	938	247	AZ676392
15	56.8	1.5	879	245	AZ550718
16	56.4	1.5	476	165	BE224646
17	55.8	1.5	910	245	AZ544844
18	55.8	1.5	1159	219	CNS015XR
19	55.4	1.5	773	220	CNS01VWG
20	55.2	1.5	735	221	CNS04NSM
21	55.2	1.5	917	219	CNS017SL
22	54.2	1.5	1076	222	CNS05HXN
23	54.2	1.5	1282	80	BF264918
24	53.8	1.5	619	258	FR0047601
25	53.8	1.5	906	245	AZ547385
26	53.8	1.5	918	247	AZ677162
27	53.4	1.4	560	219	CNS00C13
28	53.4	1.4	1101	219	CNS01807
29	53	1.4	1025	219	CNS01402
30	53	1.4	1147	219	FR0048173
31	52.4	1.4	501	258	FR0048173
32	51.8	1.4	959	219	CNS00655
33	51.8	1.4	969	219	CNS00IDL
34	51.8	1.4	1101	219	CNS00HXN
35	51.6	1.4	438	174	BG226596
36	51.6	1.4	470	222	FR0018463
37	51	1.4	911	247	AZ668666
38	51	1.4	1042	219	CNS0148K
39	51	1.4	1184	221	CNS04PAP
40	50.8	1.4	421	127	BE580431
41	50.8	1.4	460	114	BG224667
42	50.8	1.4	888	245	AZ549065
43	50.8	1.4	1101	219	CNS0039G
44	50.6	1.4	539	114	AM329517
45	50.6	1.4	1101	219	CNS00EVL

## ALIGNMENTS

RESULT 1  
 LOCUS AZ548467 908 bp DNA  
 DEFINITION ENTEK30TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.  
 ACCESSION AZ548467  
 VERSION AZ548467.1 GI:11172102  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica.  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 908)  
 AUTHORS Loftus, B., Van Aken, S. and Fraser, C.  
 TITLE Determination of clone end sequences from Entamoeba histolytica  
 HMI:IMSS sheared DNA library  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Brendan J Loftus

Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: b.loftus@tigr.org  
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library  
 Seq primer: M13-Reverse  
 Class: Shotgun  
 High quality sequence start: 17  
 High quality sequence stop: 828.  
 Location/Qualifiers  
 1..908  
 /organism="Entamoeba histolytica"  
 /strain="HMI:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_1ib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: PHOS1; Site1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, I.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT 434 a 16 c 289 g 169 t  
 ORIGIN

Query Match 2.0%; Score 72.2; DB 245; Length 908;  
 Best Local Similarity 44.6%; Pred. No. 2.8e-06;  
 Matches 333; Conservative 0; Mismatches 408; Indels 6; Gaps 1;

OY 195 taagatgtatattatgtgaatacaacaagcgaataaagcgtatgcgttagcatt 254  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 42 TGATGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGA 101  
 OY 255 aatgaataaagcagcgtgcgaataaagcgttcttagcgttagcattcaaaaata 314  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 102 AGATGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGA 161  
 OY 315 tgaacttactgttcaagaacacctaactctgcgaagcgtgtagcaactacat 374  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 162 TGAAGACGACGAATATGAATTTAGAAATGATGATGATGAAGAAGAAAGATGATGA 221  
 OY 375 cgatgttcaactatgcatacaacaataattagacgaatgcgcaagcgttagcgtcgt 434  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 222 AGATGATGAAGAAGATGATGAAGAAGATGATGATGATGATGATGATGATGATGAT 275  
 OY 435 tcaagcaaaagatttgaataaagcagaacatacttccaaaatccttataaata 494  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 276 TGAAGAAGAAGATGATGATGAAGAAGATGATGATGATGAAGAAGATGATGAAGA 335  
 OY 495 aatgcacacatcaatttagatgcgtatagttaaacacatcgtgattacttcgc 554  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 336 AGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGA 395  
 OY 555 tacatttaagcaaaagcagaacatcgcgcacacgttaattatgatatattacgctgc 614  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 396 TGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGA 455  
 OY 615 aatgaagcgcgcgaagtcagaagcgttgaagcgaagcatttagacaagcctaagc 674  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 456 AGAAGTGTGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGA 515  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 675 tgcgtgtgataatcaatacttaccaaaagtaaacagcttcaaacctgaact 734  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db	516	AGATGATGATGAAGAACAGATGACATGATGATGAAGAAGAAAGATGATGATGAAGAA	575
Qy	735	aacgaagtcgcgaaaaaagcaltgagtcgagtcgaagctgcgttactcctcaaaagttga	794
Db	576	AGATGAAGAAGAACAGAACATGATGAGAGAGTGATGTAAGAAAGAAAGAACATGATGATGAAGA	635
Qy	795	aagctgaagtcgcatcaccctcaaaaacgaagctgttgaattaaacgcgcagtcagtgaa	854
Db	636	CGACGAATATGATTTAGAACATGATGACAGACAGTGTGAAGATGAGATGATGAAGAAAGA	695
Qy	855	cysaacactaaattacaacttcacgtcgtctgcgaatgaagatacagtaaacgttaaac	914
Db	696	TGATGAAGAAGAACAGAACATGATGATGATGATGATGATGAAGAAGATGATGATGAAGATGATGA	755
Qy	915	tttacgtatctataaagtygaagctga	941
Db	756	AGAAGATGAAGATGATGACATGATGA	782
RESULT	2		
LOCUS	AZ531291/c		
DEFINITION	AZ531291	877 bp	DNA
ACCESSION	ENTB034FR	Entamoeba histolytica	Sheared DNA
VERSION	AZ531291		Entamoeba histolytica genomic C, DNA sequence.
KEYWORDS	AZ531291.1	GI:11085838	GSS.
SOURCE			
ORGANISM	Entamoeba histolytica.		
REFERENCE	Entamoeba histolytica.		
AUTHORS	Eukaryota; Entamoebidae; Entamoeba.		
TITLE	1 (bases 1 to 877)		
JOURNAL	Loftus, B., Van Aken, S. and Fraser, C.		
COMMENT	Determination of clone end sequences from Entamoeba histolytica HMI:IMSS sheared DNA library		
	Unpublished (2000)		
	Contact: Brendan J Loftus		
	Department of Eukaryotic Genomics		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel: 301 838 0200		
	Fax: 301 838 0208		
	Email: b1loftus@tigr.org		
	Clones are derived from the Entamoeba histolytica HMI:IMSS sheared		
	DNA library		
	Seq primer: M3-Reverse		
	Class: shotgun		
	High quality sequence start: 22		
	High quality sequence stop: 829.		
FEATURES			
source	1..877		
	Location/Qualifiers		

BASE COUNT	156 a	268 c	29 g	424 t
ORIGIN				

Query Match	1.8%;	Score 68.2;	DB 245;	Length 877;
Best Local Similarity	45.0%;	Pred. No. 2.4e-05;		

	Matches	299;	Conservative	0;	Mismatches	363;	Indels	3;	Gaps	1
OY	108	agtagtaagccaagcaanaagcacagcttcaaaaaagacatactactactacagccatacagct	167							
Db	825	ATTGAAACACGACGATGCAAGACGATATCAATTACAAAGCAAGCAAGATGATGATGCA	766							
OY	168	aacggaacactggttgaaattccccaacattacagatgtatagtcttgatatacaacaagcgaa	227							
Db	755	TGCAACAAACAGCAAGATGATGAAGATGATGATATATCAAGCAAGATGATGAAGCAAGATGCA	706							
OY	228	aaaagcatcaccgtagtagcgtagcatatagtgatataagcaggtggcgcaaaaaagacgc	287							
Db	705	TGATCAAGACAGATGATGATGATGATGAAGCAACATGATACAGAGATGATGAAGCAAGCAAGCA	646							
OY	288	ttacttaqctgattttacaaaagaatalgaaacttcgcttttcaaaagcaaaccttaatc	347							
Db	645	TGATGAAGTATGAAGATGATGAAGACGATGGAACGACGCAATGTGAATTGAAGATGATGCA	586							
OY	348	tggcgaagctcggttagcaaaacttacatcgatgcttacaacactgtcaacaabaattagaaga	407							
Db	585	TGATCAACAAACAGCAAGATGATGAAGATGATGATGAGCGATGGAAGCAAGCAATATGCA	526							
OY	408	aatggcccaagcgctagagcgctgcgtgttccaagcaaaaaggtttagaaaaaagcagaaacata	467							
Db	525	ATTAG---AAGTGTGTGATGATGATGAAGGAAATAATGATGTATGAGAGCAAGCAAGATGA	469							
OY	468	ctatcacacaacttcccttacttgaaattaaaaactgcgacagctacattttagatcgcgtatag	527							
Db	468	TGATGAAGTATGATGATGAAGAAAGATGCAAGATGATGAAGCAAGATGATGAAGATGA	409							
OY	528	taaaaacaactcgltgatttacttcgctctacattttaaagcaaaaagcaaaagcaactcgcgga	587							
Db	408	TGATCAACAAACAGCAAGATGATGAAGATGATGATGAAGCGATGGAAGCAAGCAATATGCA	349							
OY	588	cagcttaatttatgattatccgtgtgcgaatgaaagcgcgcggaagatgaacaagcgcgttgaa	647							
Db	348	ATTGACAAATGATGTGATGGAAGAAAGAAAGATGATGATGCAAGCAAGCAAGATGATGATGA	289							
OY	648	agcaggcaattttagaacaagaactaaagctgcgtgttgatcaaatccaatccaacttaccaaa	707							
Db	288	AGATGATGATGATGAAGCAAGCAAGATGGAAGATGATGATGAAGCAAGCAAGATGATGATGA	229							
OY	708	agtaacagatgcttccaanaactgaaactaacagaaagtagcgaaaaaagcatltagatgcaga	767							
Db	228	AGAAACAAATGATGATGATGAAGCAAGATGATGATGAAGCAAGCAAGATGATGATGAAGCAAGATGATGA	169							
OY	768	tgaag 772								
Db	168	AGAG 164								

RESULT	3
LOCUS	A2551092
DEFINITION	A2551092 912 bp DNA GSS 14-NOV-2000 ENTFU22Tf Entamoeba histolytica Sheared DNA Entamoeba histolyticae
ACCESSION	A2551092
VERSION	A2551092.1 GI:11176393
KEYWORDS	GSS.
SOURCE	Entamoeba histolytica.
ORGANISM	Entamoeba histolytica. Eukaryota; Entamoebidae; Entamoeba, 1 (bases 1 to 912)
REFERENCE	Loftus,B., Van Aken,S. and Fraser,C. Determination of clone end sequences from Entamoeba histolytica
AUTHORS	HMI:HMS sheared DNA library
TITLE	Unpublished (2000)
JOURNAL	Contact: Brendan J Loftus
COMMENT	

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208



Oy	240	tgaatgcggtgatcatctagtgaataagaacgagttggccgcgaaaaaaagacgcttacttgctcga	299
Db	694	TGATGAAGAAGATATGATGTGAAGAAGATGTGTGAAGAAGATGTGAAGAAGATGA	635
Oy	300	tttcacaaaagaatatgaaccttaacgttttccaagcaaacccctaattctcgcgaaagctcg	359
Db	634	TGATGAAGAAGATGCATCATGATCATGAAGTGTGCATGTGTGAAGAAGATGATGATGCACGA	575
Oy	360	tgtagcaacthaacatcgatgctttaacaactatgcaacaaaattagacgaatgycgccaaaga	419
Db	574	TGAAGAAGACCGATGAMAGAACGATGCAAGAACGATGCAAGAACGATGCAAGAACGACA	515
Oy	420	gcttagagcctgcgtgttcaagcaaaagatttagaaaaagcnapaacaatactatcacaaaat	479
Db	514	TGAAGAAGACGATGAMAGAACGATGCAAGAACGATGCAAGAACGATGCAAGAACGA	455
Oy	480	tccttatgaaatlaaactcgcaacgcatctttagatcgcgtlatatgytaaaacaactcg	539
Db	454	CGATGATGTAATTTCGAATTAGAAGCCATGTTGAAGAAGAACGATGATGAAGAAGACGATGA	395
Oy	540	tgatctacttcogctcatacatttaaagcaaaagcacaagaacttcgcgacagcttaatta	599
Db	394	TGAAGAAGACCGATGATGAMAGAACGATGATGCAAGAACGATGATGCAAGAACGATGA	335
Oy	600	tgatatataccgtttgcacatgaaagccgcgcaagafacaaagocgttgyaaagcgcgcaattt	659
Db	334	TGAAGAAGACCGATGATGAMAGAACGATGCTGTGAAGAAGAACGATGATGAAGAAGACGATGA	275
Oy	660	agacaaagctaaagctgctgttgtatcaaatcaatcaactacttacaaaagtacacagatgc	719
Db	274	TGAAGAAGACCGATGATGAMAGAACGATGATGCAAGAACGATGATGCAAGAACGATGA	215
Oy	720	tttcaaaactcgaactaacacgaagtagcgaaaaaagcattagatgcagatgaaagctgcgct	779
Db	214	TGAAGAAGACCGATGATGAMAGAACGATGATGCAAGAACGATGATGCAAGAACGATGA	155
Oy	780	tactccaaaagttgaagtgtaagtgcgarttaaccactcaaaacaaagctgttga	833
Db	154	TGAAGAAGATGATGAMAGAACGATGATGCAAGAACGATGCAAGAACGATGCAAGAACGA	101
RESULT	5	A2551618	843 bp DNA GSS 14-NOV-2000
LOCUS		A2551618/c	
DEFINITION		KMTDVS4fR Entamoeba histolytica Sheared DNA	Entamoeba histolytica genomic, DNA sequence.
ACCESSION		A2551618	
VERSION		A2551618.1	GI:11176919
KEYWORDS		GSS.	
SOURCE		Entamoeba histolytica.	
ORGANISM		Entamoeba histolytica.	
REFERENCE		Eukaryota; Entamoebidae; Entamoeba.	
AUTHORS		1 (bases 1 to 843)	
TITLE		Loftus,B., Van Aken,S. and Fraser,C.	
JOURNAL		Determination of clone end sequences from Entamoeba histolytica	
COMMENT		HM1:IMSS sheared DNA library	
		Unpublished (2000)	
		Contact: Brendan J Loftus	
		Department of Eukaryotic Genomics	
		The Institute for Genomic Research	
		9712 Medical Center Dr., Rockville, MD 20850, USA	
		Tel.: 301 838 0200	
		Fax: 301 838 0208	
		Email: bjl@loftusellgr.org	
		Clones are derived from the Entamoeba histolytica HM1:IMSS sheared	
		DNA library	
		Seq primer: M13-Reverse	
		Class: shotgun	
		High quality sequence start: 39	
		High quality sequence stop: 836.	
FEATURES		Location/Qualifiers	
source		1..843	

Query Match	1.8%	Score 66:	DB 245:	Length 843:
Best Local Similarity	43.7%	0	Fred. No. 7.7e-05:	
Matches 340:	Conservative	1	Mismatches 435:	Indels 3; Gaps 1;
<p>BASE COUNT 155 a 266 c 18 g 404 t</p> <p>ORIGIN</p>				
OY	119	aagcaaaagcaagcttccaaaaagacatactactactacagccatcacagtaacggaacatg	178	
DB	816	ATGAAGAAGATGATGATGATGAAGAAAGATGATGATGATGAAGAAAGATGATGAAGAAAGATGATG	757	
OY	179	gtgaatccccaacaactaacgattatattgcttgaatcacacaagcgaaaaagatacc	238	
DB	756	AAGAAGAAAGATGATGATGATGAAGATGAAGATGAAGAAAGATGATGAAGAAAGATGATGAT	697	
OY	239	gtgatgcgtagcatcttagtgaataagcagcggtgcgcgaaagaagcgcttacttagctg	298	
DB	656	TAGAAGATGATGATGATGATGAAGAAAGAAAGATGATGATGAAGAAAGATGATGAAGAAAGATG	637	
OY	299	atttcaaaaaaanaatataaactacgttcttcaaaagcaaaccttaactctggcgaagctc	358	
DB	636	ACGAGCAATGATGATGATGATGAAGATGATGATGATGAAGAAAGAAAGATGATGATGAAGAAAG	577	
OY	359	gtgtgtagcaacttaactgccttcaactatgatacaaaaattagaagaaatgagccgaag	418	
DB	576	AAGAAGATGATGATGATGAAGATGATGAAGAAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATG	517	
OY	419	agctagagcgctgctgttcaagcaaaaagattagaagaagcagaacaactactaccaaaa	478	
DB	516	ATGAAGATGATGATGATGAAGAAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAG	457	
OY	479	ttccttatgaaatataaactgcgcaagtcattttgatacgcgtatattgtaaaaaaactc	538	
DB	456	ACGAATGATGATGATGATGAAGAAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAG	397	
OY	539	gtgatttacttcgcctcatctttaaagcaaaaagcaagaactcttcgacagcttaattt	598	
DB	396	ATGATGATGATGATGATGAAGAAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATG	340	
OY	559	atgatataacgcttgcaatgaaagcgcgcgagaagtaacaagcgctgtgaagcaggaactt	658	
DB	339	AAGATGATGATGATGAAGAAAGAAAGATGATGAAGAAAGAAAGATGATGAAGAAAGAAAGATGATG	280	
OY	659	tagaacaagctaaagctgctgttgttcaatcaatcaactactaccaaaagaataacagtg	718	
DB	279	AAGAATGATGATGAAGAAAGAAAGATGATGAAGAAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATG	220	
OY	719	cttcaacaactgaacatacagaagtagcgaaaaagcaacttagatcagatgaagcgtgcgc	778	
DB	219	ATGAAGAAGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATGATG	160	
OY	779	ttactccaaggttgaagaagtgtgaagctgcatatacaactccaacaagaagcgtgttgaattaa	838	
DB	159	ATGATGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAGATGATGAAGAAAG	100	

QY	229	aaacgataccgtgatagcggtgatgacattagatgaataagcaggtggtgcgcaaaaaaacgct	288
Db	570	GACGAAGAAGATGGAAGACGAAGAAATGTAAGACGAAGATGAAGACGAACATGATGAA	511
OY	289	tacttagctgatttccaaagaataatgaacttaacgtttccaaagcaaaccttaactct	348
Db	510	GACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGAA	451
OY	349	ggcggaagctcgtgtgacgaactacacgcgatgcttaacaactatgcaacaatlaagcga	408
Db	450	GACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGAA	391
OY	409	atgcgcgaagcgtagagagctgctgtgtccaaagcaaaagattgaaaaaagcagacaatc	468
Db	390	GACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGAA	331
OY	469	tatcacaatctccattgaataatgaactcgcagacatctttagctgcgtatagct	528
Db	330	GACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGAA	271
OY	529	aaaacaactcgtgatttacttcgcctcctacattgaagcaaaaagcacaagacatccgac	588
Db	270	GACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGAA	211
OY	589	agcctaattatgatataatccgtgtgcaatgaagcgcgcgaagatcaagaagcgtgtgaa	648
Db	210	GACGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	151
OY	649	gcagcgaatctagacaaagcctaaagctgctgtgtgca	686
Db	150	GATGATGAAGACGAAGATGATGAAGACGAAGATGATGA	113
RESULT	7		
LOCUS	AZ546009	849 bp	DNA
DEFINITION	ENTFM33TF Entamoeba histolytica Sheared DNA	Entamoeba histolytica genomic, DNA sequence.	14-NOV-2000
ACCESSION	AZ546009		
VERSION	AZ546009.1	GI:11167130	
KEYWORDS	GSS.		
SOURCE	Entamoeba histolytica.		
ORGANISM	Entamoeba histolytica		
REFERENCE	Entamoeba histolytica		
AUTHORS	Eukaryota; Entamoebidae; Entamoeba.		
TITLE	1 (bases 1 to 849)		
JOURNAL	Loftus B., Van Aken S. and Fraser C.		
COMMENT	Determination of clone end sequences from Entamoeba histolytica HMI:IMSS sheared DNA library		
	Unpublished (2000)		
	Contact: Brendan J Loftus		
	Department of Eukaryotic Genomics		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel: 301 838 0200		
	Fax: 301 838 0208		
	Email: bjloftus@tigr.org		
	Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library		
	Seq primer: M13-Forward		
	Class: Shotgun		
	High quality sequence start: 26		
	High quality sequence stop: 796.		
FEATURES	Location/Qualifiers		
source	1..849		
	/organism="Entamoeba histolytica"		
	/strain="HMI:IMSS"		
	/db_xref="taxon:5759"		
	/clone_id="Entamoeba histolytica Sheared DNA"		
	/note="Vector: pHOSt1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.		
	Genomic DNA isolated from broth cultures of E. histolytica C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a		

method for isolate identification. Exp. Parasitol. 77:450-). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

BASE COUNT 199 a 236 c 43 g 371 t  
ORIGIN

Query Match 1.7%; Score 62; DB 245; Length 849;  
Best Local Similarity 44.2%; Pred. No. 0.00066;  
Matches 303; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

```
Oy 93 tacagatgtagcaacagtagtaagcacaagaacacagcttcaaaaagctactatc 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 758 TGCCGAGAGAGATTCCTTGGACTTACATGAAATGAGAAATGAAACATGATCCAGTT 699

Oy 153 ttacagcacaagtaacggaactcgtgaattcccaacattacagatglatcgtga 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 698 AGAAGACATGATTTAGAGAGAGATGATGATGATTTGAAATTGAGAGACGAGATGA 639

Oy 213 atacacaagaacgaaaaacgataccgtagtcggtgacatagtgaaacagcgtg 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 638 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 579

Oy 273 cgcgaataaagcgccttacttagctgatttcaaaaagatatgaacttgcgtttcaa 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 578 ACAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 519

Oy 333 aacaaacccataatcgtgcgaagcgtgtagcaacttaccatcgtgcttacaactatgc 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 518 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 459

Oy 333 aacaaatagaacgaatcgcgaagcgtgtagcgtgctgcttcaagaacaaattaga 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 CGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 399

Oy 453 aaaaacgaacacatctatcacaataatccttatgaataaactcgcacgcatlitt 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 339

Oy 513 agatgcgtatatagttaaaacaactcgtgatttactcgtcttacaattaaagcaaacg 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 279

Oy 573 acaagaactt--cgcgaacgcttaattatgataltaaccgttgaatgaacgcgcgga 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 TGATGAATTTGAAATTGAAAGATGAAGATGAAAGATGATGATGATGATGATGATGA 219

Oy 630 agtacaagaacgtcgtgaagcagcgcaatttagacaagaagcgtgaagcgtgctgcatcaat 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 159

Oy 630 caatcaatacttacaagaatgaatgcttcaaaaactgaacacgaagaatgacgga 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 99

Oy 750 aaaagcattagatgcagatgaacgtg 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 73
```

RESULT 8  
LOCUS AZ683582 891 bp DNA GSS 14-DEC-2000  
DEFINITION ENRK47TR Entamoeba histolytica Sheared DNA Entamoeba histolytica  
ACCESSION AZ683582  
VERSION AZ683582.1 GI:11820728  
KEYWORDS GSS.

## SOURCE

Entamoeba histolytica.

## ORGANISM

Entamoeba histolytica

Eukaryota; Entamoebidae; Entamoeba.

## REFERENCE

1 (bases 1 to 891)

Loftus, B., Van Aken, S. and Fraser, C.

Determination of clone end sequences from Entamoeba histolytica

HMI:IMSS sheared DNA library

Unpublished (2000)

JOURNAL

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: bjoftus@eigf.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Reverse

Class: Shotgun

High quality sequence start: 16

High quality sequence stop: 694.

Location/Qualifiers

1..891

/organism="Entamoeba histolytica"

/strain="HMI:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: pHS1. Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450-). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barrell, Oxford University Press, 1999).

## FEATURES

source

## BASE COUNT

400 a 55 c 291 g 145 t

## ORIGIN

Query Match 1.7%; Score 61; DB 248; Length 891;  
Best Local Similarity 43.5%; Pred. No. 0.0011;  
Matches 277; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

```
Oy 197 acgatgtatagtcgtaatacaacaagcgaacaaacgataccgtagtcggtagcatag 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 ATGATGAATTTGAAATTGAAAGATGAAGATGATGATGATGATGATGATGATGA 122

Oy 257 tgaataaagcagtggtgcggaataaagcgttacttagctgatttacaataaagatag 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 ATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 182

Oy 317 aaactttagcttcaacaaccccttaactcgtgggaagcgtcgtgtgaacttacaatgc 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 242

Oy 377 atgcttacaactatgcaacaaatagagcgaatcgcgaagaagctagaggtcgtgttc 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 ATGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 302

Oy 437 aagcaaaagattgaataaagcagaacatactatcacaataatcccttatgaaatlaaaa 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 362

Oy 497 ctgcagcgtatcttagatcgcgtatagtgttaaacacactcgtgatttacttcgtccta 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 ATGATGAATTTGAAATTGAAAGATGAAGATGAAGATGATGATGATGATGATGATGA 422

Oy 557 catttaagcaaaagcacaagaacttcgcgaagcttcaattatgataltaaccgttgcaa 616
```



```

Db      423  AAGAAGACGATGATGAAGACGATGATGAAGACGATGATGAAGACGATGATG 482
Qy      617  tgaagcgcggaagtcacaagacgctgtgaaagcggaattagacaagctaaagctg 676
Db      483  AAGAAGACGATGATGAAGACGATGATGAAGACGATGATGAAGACGATGATG 542
Qy      677  cgttgatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 736
Db      543  AAGAAGACGATGATGAAGACGATGATGAAGACGATGATGAAGACGATGATG 602
Qy      737  cagaagtcgcaaaaaagcattagatgcagatgaagctgcgtcttaccacaaagt 796
Db      603  AAGAAGACGATGATGAAGACGATGATGAAGACGATGATGAAGACGATGATG 662
Qy      797  gtgtaatgtcggaattacactcaaaacaaagctgttga 833
Db      663  ACGACGATGAAGACGATGAAGACGATGAAGACGATGAAGACGATGAAGACGAT 699

```

```

RESULT  9
LOCUS   A2549980      900 bp      DNA      GSS      14-NOV-2000
DEFINITION  ENTDD94TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
ACCESSION  A2549980
VERSION    A2549980.1 GI:11175122
KEYWORDS   GSS.
SOURCE     Entamoeba histolytica.
            Entamoeba histolytica.
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE  1 (bases 1 to 900)
            Loftus, B., Van Aken, S. and Fraser, C.
            Determination of clone end sequences from Entamoeba histolytica
            HML:IMSS sheared DNA library
            Unpublished (2000)
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: b.loftus@tigr.org
            Clones are derived from the Entamoeba histolytica HML:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: Shotgun
            High quality sequence start: 20
            High quality sequence stop: 890.
            Location/Qualifiers
                1..900
                /organism="Entamoeba histolytica"
                /strain="HML:IMSS"
                /db_xref="taxon:5759"
                /clone_lib="Entamoeba histolytica Sheared DNA"
                /note="Vector: pHOSt1; Site1; Bst I; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from broth cultures of E. histolytica
                using a method described by Clark and Diamond (Clark,
                C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                method for isolate identification. Exp. Parasitol.
                77:450.). The DNA was mechanically sheared to give a
                tight size distribution (~2 Kb). The v + i method used for
                the library construction is described in detail in Smilh,
                H.O. and Venter, J.C. (Making small insert libraries for
                whole genome shotgun sequencing projects. In Genome
                Sequencing: A Practical Approach, eds. M. Vaudin and B.
                Barrell, Oxford university Press, 1999)."
```

```

BASE COUNT      376 a      63 c      235 g      226 t
ORIGIN

```

Query Match

1.6%; Score 60; DB 245; Length 900;

Best Local Similarity 44.8%; Pred. No. 0.002;  
Matches 231; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

```

Qy      260  ataagaagctgctgcgcaaaaaagacgcttaactagctgattacaagaatgatga 319
Db      254  ATATGACTGATGATTTGAAGAACATGAAGATTTAGATGACGAGAGATTTGACTTGA 313
Qy      320  cttagctttcaagaaccccaaatctgcggaactcgtgtgagaacttacatgatg 379
Db      314  ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGAT 373
Qy      380  cttaacactgcgaacaaattagacgaatgcgcaagaagctgaaggctgtcttca 439
Db      374  ACGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 433
Qy      440  caaagaatttgaaaaagacgaacaaatcaatcaacaaatctttagaattaaact 499
Db      434  AAGAAATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATG 493
Qy      500  gcacagtcatttagatgcgtatgtgtaaaacaactcgtgatttaacttcgtct 559
Db      494  ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGAT 553
Qy      560  ttaagcaaaagcacagaacttcgcgacagcttaattatgattaccgttcgaatg 619
Db      554  ATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 613
Qy      620  aagcgcggaagtcacaagacgctgtgaaagcagcaatttagacaagctgaagct 679
Db      614  AAGAAATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATG 673
Qy      680  ttgatcaaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 729
Db      674  ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGAT 733
Qy      740  aagtagcaaaaaagcattagatgcagatgaagctg 775
Db      734  ATGAAGACGATGATGAAGACGATGATGAAGACGATG 769

```

```

RESULT  10
LOCUS   CNS0134P      997 bp      DNA      GSS      26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
ACCESSION  BACN09C07 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
VERSION    AL102403
KEYWORDS   GSS.
SOURCE     fruit fly.
            Plasmid Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 997)
            Genoscope.
            Direct Submission
            Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billard at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBelobAC11.
            Location/Qualifiers
                1..997
                /organism="Drosophila melanogaster"
                /plasmid="pBelobAC11"
                /db_xref="taxon:7227"
```

```

FEATURES
    source

```

BASE COUNT	162 a	115 c	79 g	517 t	124 others
ORIGIN	/clone_lib="DrosBAC" /clone="BACN09C07" /note="end : Sp6"				

Query Match	1.6%;	Score 59.8;	DB 219;	Length 997;
Best Local Similarity	37.5%;	Pred. No. 0.0022;		
Matches 245;	Conservative 43;	Mismatches 362;	Indels 3;	Gaps 1.

[illegible][illegible]

Qy	677	ctgttgatcaaatcaatcaatacttaccaaagaatcacagatgcttccaactgaactaa	736
Db	416	wtgmaaaagaaaaataatataatwtwmaataatagacagratatataatmgcaataa	3573
Qy	737	cagaagtagcgaanaaagcattagatgcagatgaagctgcttactccaagaattgaa	736
Db	356	atgcgaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaaata	3573

[illegible][illegible]

DB 119 TTACCAATVTAAGGMAATRAAAAAAAAAAATTAATAAAGTTAAACAAAAAGAAACGGTAAWTTAA 60

QY 1037 agtaaaagtaagtaagtaaaagtaataaacaacaaaatggcaagaattttaa 1089

DB 59 ACCGGAACACAGTTGTAAAAAATGAAAAAATAATTATTGMAATGMAATTAAM 7

RESULT	11			
LOCUS	FR0048073			
DEFINITION	FR0048073	494 bp	DNA	05-JAN-2001
ACCESSION	Fugu rubripes GSS sequence, clone 263K15bD8,			
VERSION	AL444858			
KEYWORDS	AL444858.1	GI:12052694		
SOURCE	GSS: genome survey sequence.			
ORGANISM	Takifugu rubripes.			
	Takifugu rubripes			

REFERENCE	1 (bases 1 to 49)
AUTHORS	Clark, M.S.
TITLE	Direct Submission

**JOURNAL** Submitted (29-SEP-2000) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10 1SB, UK Email: biiohelphgmp.mrc.ac.uk  
**COMMENT** Vector: pBluescript II KS

**FEATURES**

- One pass dye-terminator sequencing of BAC (pBelobAC11) cloned genomic sequence
- The BACs can be obtained from <http://www.incyte.com>.
- location/qualifiers

```

BASE COUNT      128 a      163 c      45 g      158 t
ORIGIN
    /organism="Rakifugu rubripes"
    /db_xref="taxon:31033"
    /clone="26K15D8"
    /clone_lib="BAC 263K15"

```

Query Match	1.6%	Score 58.8	DB 256	length 494
Best Local Similarity	48.0%	Pred. No. 0.0032		
Matches 168; Conservative	0;	Mismatches 182;	Indels 0;	Gaps 0

[illegible]

QY	1379	ttgtatcaataatgtttaaacataactcgtgttacaagcagaagcttaacgaatagtgt	1438
Db	265	CTACTACTACTACTACTACTACGACGACGACACTACTACTCTACTACTACTACTA	324
QY	1439	tcactgaaacgcatagcaccacactgtgtgtacgcgcctctactacttaagtgttaca	1498
	225	caatgaaatgct	284

[illegible]

RESULT	12
A2550256/c	
LOCUS	
DEFINITION	
<p>A2550256 905 bp DNA</p> <p>ENTREV58fR Entamoeba histolytica Sheared DNA GSS 14-NOV-2000</p> <p>Genomic DNA sequencing</p>	

ACCESSION	GenBank, vna sequence.
VERSION	AZ550256
KEYWORDS	AZ550256.1 GI:1179557
SOURCE	GSS.
ORGANISM	Entamoeba histolytica.
REFERENCE	Entamoeba histolytica
AUTHORS	Eukaryota; Entamoebidae; Entamoeba.
TITLE	1 (bases 1 to 905)
JOURNAL	Loftus, B., Van Aken, S. and Fraser, C.
COMMENT	Determination of clone end sequences from Entamoeba histolytica
	HMI:IMSS sheared DNA library
	Unpublished (2000)
	Contact: Brendan J Loftus

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: [bjloftus@igr.org](mailto:bjloftus@igr.org)  
Clones are derived from the *Entamoeba histolytica* HMI:INSS shared







**THIS PAGE BLANK (USPTO)**